Xiaoxin Zhou 1003549742
INF 2178 A4

Introduction:
The dataset collection provides a detailed exploration of brain aging and the progression of Alzheimer's Disease (AD) within a diverse population. The cross-sectional component encompasses MRI scans from 294 individuals, ranging in age from 60 to 98. This segment is notably varied, documenting the brain structures of young, middle-aged, nondemented, and demented older adults. It notably includes data from 100 older adults diagnosed with very mild to moderate Alzheimer's Disease. For this project, mixed-effects ANOVA and statistical power analysis will be utilized. The research question: How does the mean score on a cognitive test change over time, and does this change differ between groups of individuals classified as nondemented, demented, and converted?

Data Information

	Unnamed: 0	Visit	MR Delay	Age	EDUC	SES	MMSE	CDR	eTIV	nWBV	ASF
count	294.0	294.0	294.0	294.0	294.0	279.0	293.0	294.0	294.0	294.0	294.0
mean	190.4	1.4	349.7	76.4	14.5	2.49	27.25	0.3	1478.8	0.73	1.2
std	106.6	0.5	400.7	7.6	2.8	1.12	3.4	0.38	176.5	0.03	0.1
min	0.0	1.0	0.0	60.0	6.0	1.0	15.0	0.0	1106.0	0.64	0.8
25%	99.0	1.0	0.0	71.0	12.0	2.0	26.0	0.0	1347.2	0.7	1.1
50%	195.5	1.0	0.0	76.0	14.5	2.0	29.0	0.0	1461.5	0.73	1.2
75%	282.7	2.0	671.5	81.0	16.0	3.0	30.0	0.5	1569.0	0.7	1.3
max	371.0	2.0	1707.0	98.0	23.0	5.0	30.0	2.0	2004.0	0.8	1.5

Table 1 provides a summary of the dataset, revealing that the majority of participants are seniors, as indicated by the age feature. Drop

Preprocess:
Rename M/F to a reasonable name as Gender, set 'F' with 0 and 'M' with 1 in the Gender column, and replace 'Nondemented' with 0 and 'Demented' with 1, 'Converted' with 2 in the 'Group' column. Dropping 'Unnamed: 0', 'Subject ID', 'MRI ID', 'Hand' columns.

Nondemented: From Figure 1, this group has a relatively wide age distribution, suggesting it includes a broad range of individuals from the younger to the older end of the spectrum. The median age is roughly in the mid-70s, and there are numerous data points spread above and below the median, indicating variability in the ages of nondemented individuals.

Demented: Similar to the Nondemented group, Demented individuals also exhibit a wide range of ages but with the median age slightly lower than that of the Nondemented group. There is a notable presence of outliers on the higher age end.

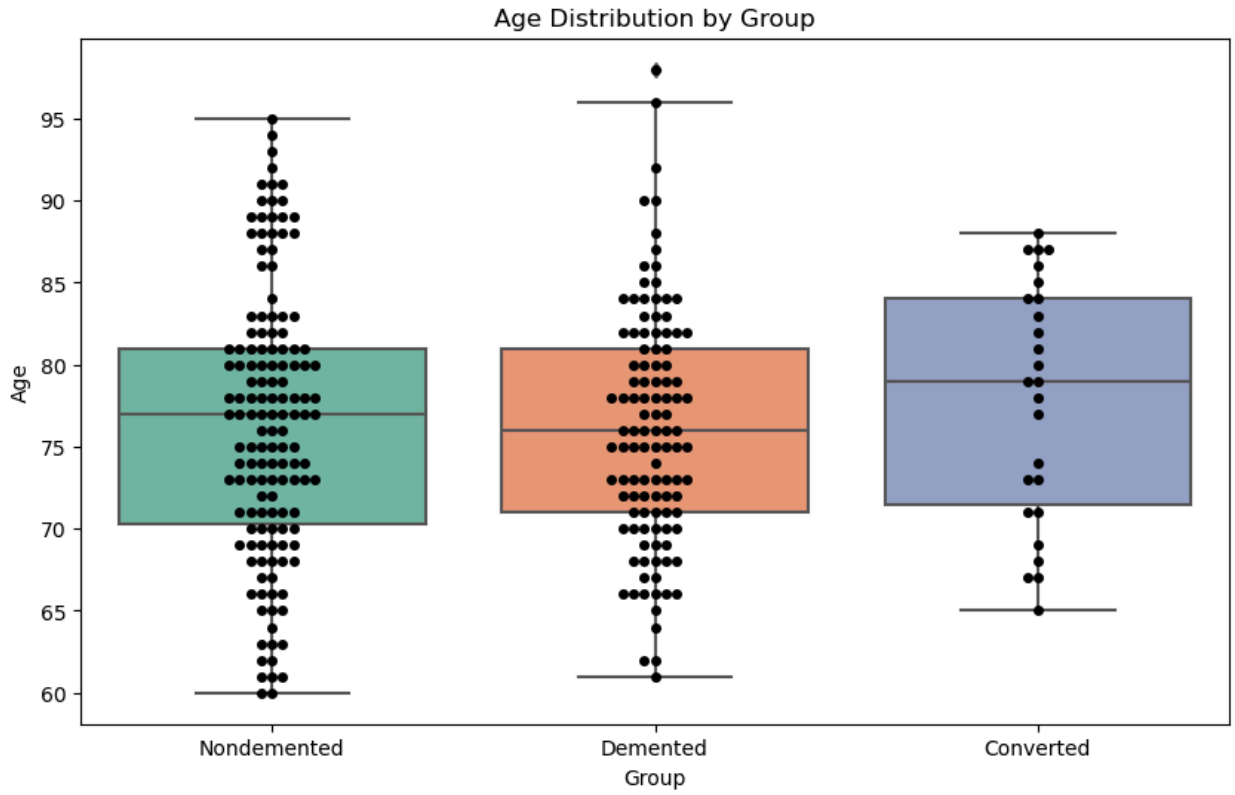

Figure 1 A visual representation of the age distribution among three distinct groups, presumably categorized based on their Alzheimer's Disease status: Nondemented, Demented, and Converted.

	Group	Visit	MR Delay	Gender	Age	EDUC	SES	MMSE	CDR	eTIV	nWBV	ASF
Group	1.000000	-0.011374	0.005043	0.137968	0.023072	-0.108774	-0.019625	-0.357346	0.531388	-0.038779	-0.203095	0.026104
Visit	-0.011374	1.000000	0.890511	0.000314	0.125618	0.012901	-0.001734	-0.091375	0.102667	0.022780	-0.125161	-0.021329
MR Delay	0.005043	0.890511	1.000000	0.003812	0.092857	0.047935	0.003741	-0.006965	0.041467	0.037858	-0.082946	-0.034321
Gender	0.137968	0.000314	0.003812	1.000000	-0.048800	0.075418	-0.032238	-0.153873	0.181896	0.568017	-0.218030	-0.559039
Age	0.023072	0.125618	0.092857	-0.048800	1.000000	-0.051711	-0.009439	-0.011998	0.008663	-0.008501	-0.549936	0.022364
EDUC	-0.108774	0.012901	0.047935	0.075418	-0.051711	1.000000	-0.731764	0.205380	-0.136017	0.245815	0.040125	-0.228690
SES	-0.019625	-0.001734	0.003741	-0.032238	-0.009439	-0.731764	1.000000	-0.196523	0.091480	-0.248358	0.019837	0.239745
MMSE	-0.357346	-0.091375	-0.006965	-0.153873	-0.011998	0.205380	-0.196523	1.000000	-0.723021	0.005115	0.350983	0.006865
CDR	0.531388	0.102667	0.041467	0.181896	0.008663	-0.136017	0.091480	-0.723021	1.000000	0.068303	-0.326791	-0.078887
eTIV	-0.038779	0.022780	0.037858	0.568017	-0.008501	0.245815	-0.248358	0.005115	0.068303	1.000000	-0.197910	-0.988011
nWBV	-0.203095	-0.125161	-0.082946	-0.218030	-0.549936	0.040125	0.019837	0.350983	-0.326791	-0.197910	1.000000	0.196757
ASF	0.026104	-0.021329	-0.034321	-0.559039	0.022364	-0.228690	0.239745	0.006865	-0.078887	-0.988011	0.196757	1.000000

Table 2, Visit and MR Delay 0.89, suggests a strong positive relationship, meaning that as MR Delay increases, so does the number of visits.

EDUC and SES -0.73, as the years of education (EDUC) increase, the socioeconomic status (SES) decreases. (it is very counterintuitive, not sure the data is correct or some other thing effect this result)
MMSE and CDR -0.72, we expect cognitive function to decline as dementia severity increases.
nWBV and Age -0.54, brain volume can shrink with age!
ASF and eTIV -0.98, ASF is calculated in a way that is inversely related to eTIV, they have strong correlation.
Group and CDR 0.53, as CDR increases, indicating more severe dementia, it becomes more likely that the individual will be classified in the 'Demented' group.
Group and MMSE -0.36, A higher MMSE score, which indicates better cognitive function, is associated with a lower likelihood of being in the 'Demented'.
Group and Gender 0.14, a weak positive correlation suggests a slight association between gender and the group classification, it looks like men have a higher chance to get 'Demented'.

Two-way mixed-design ANOVA

```
=============
ANOVA SUMMARY
=============
```

Source	SS	DF1	DF2	MS	F	p-unc	np2	eps
Group	1325.778	2	134	662.889	56.575	0.000	0.458	nan
Visit	21.639	1	134	21.639	9.001	0.003	0.063	1.000
Interaction	17.716	2	134	8.858	3.685	0.028	0.052	nan

Between-subjects factor: Categorized into three levels, likely representing 'Nondemented', 'Demented', and 'Converted'.
Within-subjects factor: Representing different time points in this case I use Visit.
Group factor:
Sum of Squares: This is a measure of the total variability in the MMSE scores attributed to the differences between the groups.
F-statistic: The value of 56.575 is very high, which indicates a strong group effect.
p-value: The p-value is less than 0.001 (assuming a standard format where $p < 0.001$ is denoted as 0.000). This indicates that the differences between at least some of the groups are statistically significant.
Partial Eta Squared: This is a measure of effect size. With a value of 0.458, it indicates a large effect size, meaning that the group status accounts for a substantial amount of the variance in the MMSE scores.
Visit factor:
Sum of Squares: This reflects the variability in MMSE scores due to the different visits
F-statistic: The value of 9.001 is quite high, indicating that the time points have a significant effect on MMSE scores.
P-value: The p-value of 0.003 is below the standard threshold of 0.05, indicating that this effect is statistically significant.
Partial Eta Squared: 0.063 - A smaller effect size than for the group.

Group*Visit factor:
F-statistic: 3.685 suggests that the interaction effect is present.
p-value: 0.028 is less than 0.05, indicating that the interaction is statistically significant.
Partial Eta Squared: 0.052, although not as large as the group main effect.

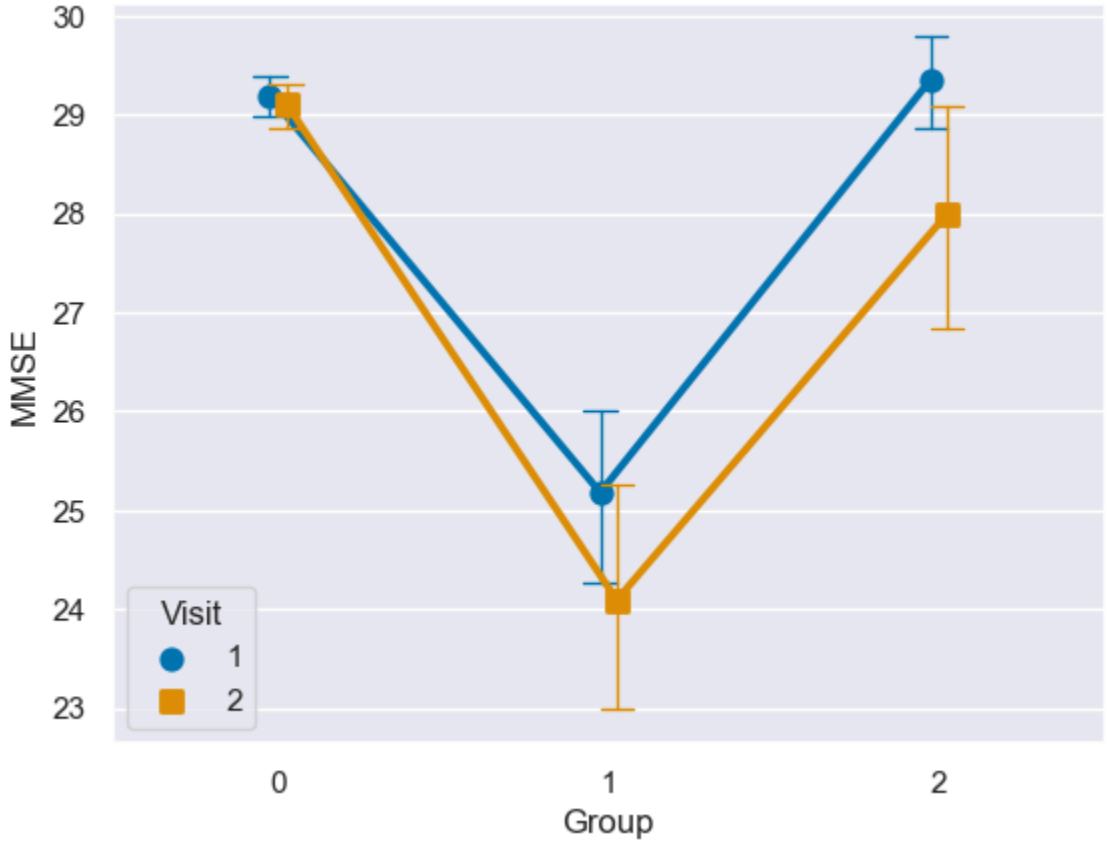

Figure 2 The lines on the plot demonstrate the pattern of change in MMSE scores across visits for each group, supporting the statistical findings that there are significant differences and interactions at play.

The main effect of Group is significant, indicating differences in cognitive scores across different groups (nondemented, demented, converted). The Visit is significant, showing that cognitive scores change over time. The interaction between Group and Visit is also significant, meaning that the change in cognitive scores over time is different across the groups. This could indicate that **demented patients' scores decrease more rapidly over time than those of nondemented patients.**

statistical power for t-tests
Power result: 0.9756621679978977
Cohen's result: 0.468411445953685
The calculated power is approximately 97.6%, which is exceptionally high. This power level means that if there is indeed an effect of the size specified by Cohen's d = 0.468, 97.6% chance of correctly rejecting the null hypothesis in a two-sided t-test for Group*Visit factor.

Conclusion:
The study underscored by significant ANOVA results and high statistical power, provides robust evidence that cognitive decline trajectories vary significantly across different stages of Alzheimer's Disease. The findings illuminate the nuanced impact of AD progression on cognitive function, with distinct patterns emerging among nondemented, demented, and converted groups. This research not only enriches our understanding of AD's impact on aging brains but also underscores the importance of targeted interventions tailored to specific AD stages.

Future research should continue to explore these patterns, potentially integrating longitudinal data to capture the full spectrum of AD progression and its implications for patient care strategies. The insights gleaned from such studies are invaluable for developing nuanced, stage-appropriate interventions that address the complex needs of individuals at different points in the AD continuum.